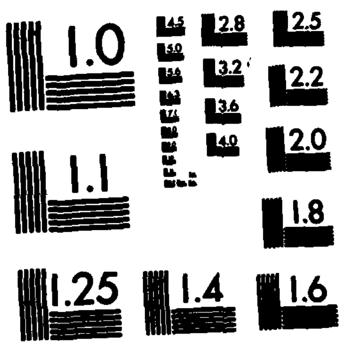
AD-A124 440 SIGNAL DETECTION IN THE PRESENCE OF WEAKLY DEPENDENT NOISE PART II ROBUST DETECTION(U) ILLINOIS UNIV AT URBANA COORDINATED SCIENCE LAB H V POOR DEC 81 R-932 1/1

UNCLASSIFIED N00014-79-C-0424 F/G 12/1 NL

END

FILMED

DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A224	3. RECIPIENT'S CATALOG NUMBER 440
4. TITLE (and Subtitle) SIGNAL DETECTION IN THE PRESENCE OF WEAKLY DEPENDENT NOISE - PART II: ROBUST DETECTION		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-932; UILU-ENG 81-2263
7. AUTHOR(s) H. Vincent Poor		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		12. REPORT DATE December 1981
		13. NUMBER OF PAGES 37
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

robust signal detection; weakly dependent noise; uncertain models; moving-average dependence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problem of designing robust systems for detecting constant signals in the presence of weakly dependent noise with uncertain statistics is considered. As in Part I of this study, which treated optimum detection in weakly dependent noise, a moving-average representation is used to model the dependence structure of the noise process, with the degree of dependence being parameterized by the averaging weights. Weak dependence is then modeled as the situation in which quantities depending to second or higher order on the averaging weights can be considered to be negligible. Uncertainty (over)

DD FORM 1 JAN 73 1473

20. (continued)

in the noise statistics is introduced within this framework by allowing a general type of uncertainty in the univariate statistics of the independent sequence that drives the moving average. To find robust detectors for signals in this type of weakly dependent noise environment, related results concerning robust location estimation in an analogous dependent situation are applied to modify a robust detection system for the corresponding independent-noise case. It is argued here that a robust detector for this dependent noise model is characterized by a least-favorable noise distribution which coincides with the distribution that is least favorable for the corresponding independent-noise case. However, the resulting detector design for dependent noise differs from that for independent noise; in particular, the robust detector for dependent noise is based on a linearly corrected version of the influence curve that defines the independent-noise robust detector. The worst-case performance of the proposed robust detector relative to that of the independent-noise robust detector is also analyzed, with the conclusion that the performance of the proposed technique is better, to first order in the averaging weights, in this respect. A modification of this robust detector is also proposed which eliminates some practical disadvantages of this system while retaining equivalent performance to first order. The specific situation of contaminated Gaussian noise is treated in order to illustrate the analysis.

UILU-ENG-81-2263

SIGNAL DETECTION IN THE PRESENCE OF WEAKLY DEPENDENT NOISE - PART II: ROBUST DETECTION

by

H. Vincent Poor

This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy and U.S. Air Force) under Contract N00014-79-C-0424.

FOREWORD

This report is a preprint of a paper with the same title which is scheduled to appear in the September 1982, issue of the IEEE Transactions on Information Theory (vol. IT-28). This is the second part of a two-part study, the first part of which is published as CSL Report No. R-931.

I. Introduction and Preliminaries

As in Part I of this study [1], we consider the signal detection problem described by the following pair of statistical hypotheses concerning a random process $\{X_i;\ i = 1,2,\ldots,\infty\}$ from which we have a sequence $\{x_i;\ i = 1,2,\ldots,n\}$ of observations:

$$H_0:\ X_i = N_i \qquad ;\ i = 1,2,\ldots,n$$

versus

$$H_1:\ X_i = N_i + \theta\ ;\ i = 1,2,\ldots,n \tag{1}$$

where $\{N_i;\ i = 1,2,\ldots,\infty\}$ is a stationary noise process and θ is a known positive signal. Part I of this study [1] considered the problem of designing asymptotically efficient detection systems for the problem of (1) in which the noise process exhibits a weak moving-average type of dependence. Specifically, [1] considered the noise model

$$N_i = \rho\ Y_{i-1} + Y_i + \rho\ Y_{i+1}\ \ ;\ \ i = 1,2,\ldots \tag{2}$$

where $\{Y_i;\ i \in Z\}$ is an independent and identically distributed (i.i.d.) noise-generating sequence with marginal probability density function (p.d.f.)p, and where ρ is a fixed parameter indexing the degree of dependence among the noise samples. It was shown in [1] that an appropriate detector for this situation with $|\rho|$ small is of the form of the corresponding optimum independent-noise ($\rho = 0$) system with an added linear correction factor depending on the degree of dependence as parameterized by ρ and on the particular noise statistics as determined by p. It was also seen in [1] that appreciable improvement over the corresponding independent-noise system can often be expected from the proposed system, particularly for channels dominated by impulsive noise.

The specification of the detection system proposed in [1] requires a complete knowledge of the statistics of the noise process $\{N_i; i = 1,2,\ldots,\infty\}$ as determined by ρ and p. An important practical modification of this problem is the situation in which the noise distribution is not known exactly but rather is known to be in some class of possible noise distributions representing uncertainty in a nominal noise model. In this situation a detector is desired whose performance does not deteriorate drastically over the class of possible noise statistics. A detection system with this general property is usually known as a robust detector. Several investigators have considered the problem of robust detection in the model of (1) for the case in which the noise sequence is independent. General results for robust hypothesis testing in a model which includes the independent-noise case of (1) have been obtained by Huber [2] and by Huber and Strassen [3] within a minimax risk (or error probability) formulation and by the author [4] within a maximin distance formulation. Martin and Schwartz [5] have considered robust detection in the model of (1) for the case in which the independent noise process is a mixture-contaminated Gaussian process. Within this context both minimax risk and maximin local power slope[1] are treated in [5]. Kassam and Thomas [8] have extended the results of [5] to solve the maximin-power-slope formulation for the case of mixture-contaminated nonGaussian noise processes, and, in an asymptotic formulation, El-Sawy and VandeLinde [9] have treated the minimax risk problem for (1) with general uncertainty classes for the marginal distribution of the independent noise sequence. A sequential version of (1) is treated in a similar context by El-Sawy and VandeLinde in [10].

---

[1]Recall that the local power slope is the criterion for designing locally most powerful detection systems (see Capon [6] or Ferguson [7]).

Considerably less work has been done on the problem of robust detection for the situation of (1) with dependent noise, although the related problem of nonparametric (or distribution-free) detection in (1) with dependent noise has been considered in some detail (see, for example, Kanefsky and Thomas [11], Gastwirth, et al. [12], Davisson, et al. [13], Kassam and Thomas [14], and Sirvanci and Wolff [15]), as has the statistically analogous problem of robust estimation of location with dependent errors (see Hoyland [16], Gastwirth and Rubin [17], Portnoy [18,19], and Koul [20]). In this paper we consider robust detection in (1) for situations in which the noise can be modeled by a weakly dependent moving-average process as described by (2) with small $|\rho|$. Since the parameter ρ can be estimated straightforwardly from an observation of $\{X_i;\ i=1,2,\ldots,\infty\}$ (see Eq. (53) of [1]) we model statistical uncertainty in the noise sequence $\{N_i;\ i=1,2,\ldots,\infty\}$ by assuming that the p.d.f. p of the noise-generating sequence $\{Y_i;\ i \in Z\}$ is not known exactly but, rather, is known only to be a member of a class $\mathcal{F}$ of symmetric probability densities. Note that, for analytical reasons, the criterion of asymptotic detection efficiency (which was applied in [1] for optimum design) is not suitable for robust design in this situation (see, for example, Martin and deMontricher [21]). Thus, we adopt a Neyman-Pearson design philosophy and, for a particular detector φ, consider the probabilities of false alarm and detection, P_F and P_D, respectively, given for $p \in \mathcal{F}$ by

$$P_F(\varphi|p) = P\{\varphi \text{ chooses } H_1 | H_0 \text{ is true and } Y_i \sim p \ \ \forall\, i \in Z\} \tag{3}$$

and

$$P_D(\varphi|p) = P\{\varphi \text{ chooses } H_1 | H_1 \text{ is true and } Y_i \sim p \ \ \forall\, i \in Z\} . \tag{4}$$

Adopting the game-theoretic philosophy of robust design (as proposed by Huber [2]) we may define an α-level robust detector for this situation as one which solves the game

$$\max_{\varphi \in \mathcal{D}} \{ \inf_{p \in \mathcal{F}} P_D(\varphi|p)\} \text{ subject to } \sup_{p \in \mathcal{F}} P_F(\varphi|p) \le \alpha \tag{5}$$

where $\mathcal{D}$ is an appropriate class of detectors. Note that (5) reduces to the traditional Neyman-Pearson design criterion if $\mathcal{F}$ consists of a single p.d.f. [7].

In this paper we seek the solution to (5) for the small-$|\rho|$ situation in (1) and (2) by restricting $\mathcal{D}$ to be the class of M-detectors as proposed by El-Sawy and VandeLinde in [9]. In so doing we are able to exploit related results obtained by Portnoy [18,19] for the problem of robust estimation of a location parameter with dependent observation errors modeled by (2) with small $|\rho|$. The class of detectors to be considered here is defined specifically in Section II and first-order (in ρ) approximations to their asymptotic false-alarm and detection probabilities are developed. The structure of M-detectors that are optimum with respect to these approximations is also presented in Section II, and it is seen that the system modification to account for weak noise dependency in M-detectors is similar to that found for the detectors considered in [1]. The first-order approximations to the error probabilities developed in Section II suggest a first-order approximation to the minimax problem of (5), and Section III considers the solution to this first-order version of (5). In particular, it is argued that a small-$|\rho|$ robust design for this problem is the small-$|\rho|$ optimum design corresponding to a least-favorable noise-generating p.d.f. Moreover, it is seen (as in the analogous situation of [18]) that this least-favorable noise-generating p.d.f. does not depend on the value of ρ and thus is the same as the p.d.f.

that is least favorable for the independent noise case. This latter result is useful because the corresponding independent-noise problem has been solved previously for a number of noise uncertainty models (see, for example, [5,8,9]). In Section IV, we consider a modification of the system developed in Section III because of qualitative disadvantages of the system of Section III. It is demonstrated analytically in Section IV that this modification is equivalent, to first order in ρ, to the detector developed in Section III, a fact which supports the use of the modification intuitively since the modified system is superior to the unmodified one. Also in Section IV, the specific example of a contaminated Gaussian noise model (as proposed by Huber [22]) is considered in some detail in order to illustrate the proposed robust design procedure. In Section V we return to the general situation to consider the degree to which the worst-case performance of the proposed robust detector outperforms the corresponding independent-noise ($\rho = 0$) robust detector. In particular, it is shown that this performance difference is of first order in ρ, a fact which indicates that the consideration of dependence is even more important in the robust design problem than in the corresponding optimum design problem of [1]. (The corresponding performance difference in [1] is of second order in ρ.) Finally, Section VI includes some further discussion of the results of this paper and of some possible extensions of these results.

II. M-Detectors and Their Performance in Weakly Dependent Noise

El-Sawy and VandeLinde [9] have solved the problem of (5) for the asymptotic ($n \to \infty$) case with independent noise ($\rho = 0$) when the class $\mathcal{D}$ is restricted to contain only detectors of a specific structure known as M-detectors. This class of detectors is based on a class of location estimates known as M-estimates introduced by Huber in [22], and certain members of this class derive their robustness properties from analogous properties for robust estimation. Robust M-estimates of location for the weak dependence model of (2) with small $|\rho|$ have been considered by Portnoy in [18] and [19], and thus it is reasonable to consider the related class of M-detectors to seek a solution to the analogous problem of (5) with $|\rho|$ small but nonzero.

We therefore restrict $\mathcal{D}$ to contain only detectors of the following form

$$\varphi_\psi(\underline{x}) = \begin{cases} 1\ ; & \text{if } \hat{\theta}_\psi(\underline{x}) > \tau \\ \gamma\ ; & \text{if } \hat{\theta}_\psi(\underline{x}) = \tau \\ 0\ ; & \text{if } \hat{\theta}_\psi(\underline{x}) < \tau \end{cases} \tag{6}$$

where $\hat{\theta}_\psi(\underline{x})$ is a solution to the equation

$$\sum_{i=1}^{n} \psi(x_i - T)\Big|_{T=\hat{\theta}_\psi(\underline{x})} = 0 \tag{7}$$

with ψ an arbitrary function, known as the influence curve of $\hat{\theta}_\psi$ (see also Hampel [23] and Huber [24]) characterizing the detector φ_ψ. Here $\varphi_\psi(\underline{x})$ denotes the probability with which we accept H_1 given that $\underline{x}$ is observed, and the threshold γ and randomization τ are chosen to yield desired false-alarm performance. A detector of the form (6) is known as an

M-detector [9] and the detection statistic $\hat{\theta}_\psi(\underline{x})$ of (7) is an M-estimate [22]. Note that the M-estimate is a generalization of the independent-noise maximum-likelihood estimate of location of $\underline{X}$ which corresponds to the particular choice of influence curve $\psi(x) = d \log[p(x)]/dx$ (see, for example, Silvey [25]). Moreover, in the independent case, the M-detector corresponding to the maximum-likelihood M-estimate is asymptotically equivalent to the likelihood ratio test of H_0 versus H_1 (see, for example, Lemma 3 of [9]). Thus, the restriction to detectors of the form of (6) is reasonable for the small-$|\rho|$ and large-n case in that it does not eliminate the optimum detector for any member of $\mathcal{F}$ from consideration when $\rho = 0$. Note further that the class of M-estimates includes the sample mean (given by $\psi(x) = x$), for which the corresponding M-detector is the linear detector, as well as the sample median (given by $\psi(x) = \mathrm{sgn}(x)$).

Note that, for many choices of the influence curve ψ, (7) will sometimes have multiple solutions; however, for analytical (and implementational) purposes, we would like to specify a particular solution to (7). Thus, if for a given $\underline{x}$ there are multiple solutions to (7) we will choose $\hat{\theta}_\psi(\underline{x})$ to be the solution closest to the sample mean $\bar{x} \triangleq \frac{1}{n} \sum_{i=1}^{n} x_i$, and, if there are two solutions equidistant from $\bar{x}$, we choose $\hat{\theta}_\psi(\underline{x})$ to be the larger of the two. Also, if (7) has no solution for a particular $\underline{x}$ we take $\hat{\theta}_\psi(\underline{x}) = 0$. With this construction of $\hat{\theta}_\psi(\underline{x})$ we may state the following result which follows from Theorem A.2 of Portnoy [18].

<u>Theorem 1</u>: Assume the model of (1) and (2). Suppose $\sigma^2 \triangleq \mathrm{Var}(Y_1) < \infty$, ψ is continuous and bounded, $E\{\psi(N_1 + \theta)\}$ is strictly increasing in a neighborhood of $\theta = 0$, and $E\{\psi(N_1)\} = 0$. Then $\hat{\theta}_\psi(\underline{X})$ converges in probability to θ under H_1 and to 0 under H_0 as $n \to \infty$.

<u>Proof</u>: For the model of (1) and (2) with $\sigma^2 < \infty$ it follows from the Chebychev inequality and from Lemma 3 of Billingsley [26 , p. 172] that $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges in probability to θ under H_1 and in probability to 0 under H_0 as $n \to \infty$. This and the fact that $\{N_i;\ i = 1,2,\ldots,\infty\}$ is 2-dependent (i.e., N_i and N_j are independent if $|i-j| > 2$) imply Theorem 1 via Theorem A.2 of [18].[2]

Thus, under the mild conditions of Theorem 1, we see that the M-detector φ_ψ of (6) provides a consistent test of H_0 versus H_1 provided that the threshold τ is between 0 and θ. That is, φ_ψ with $\tau \in (0,\theta)$ has the property that false-alarm and detection probabilities, P_D and P_F, converge to 0 and 1, respectively, as the number of samples n approaches ∞. Consistency is, of course, the very least that we should require of a detection procedure; thus in order to optimize over the class $\mathcal{B}$, it is of further interest to approximate the large-n performance of the detectors of this form for the weak dependence model of (2). For this purpose we may state the following result which follows straightforwardly from Theorems 2.1 and A.4 of [18].

<u>Theorem 2</u>: Assume the hypothesis of Theorem 1. Suppose in addition that ψ is differentiable except as a closed set D of Lebesgue measure 0, that ψ' is uniformly continuous off of D, that $0 \notin D$, that $E\{\psi'(N_i)\}$ and $E\{\psi'(Y_i)\}$ are positive, and that the characteristic function ϕ_Y of Y_i satisfies $\int u^2|\phi_Y(u)|du < \infty$. Then, with $X_i = N_i + \theta$ for $i = 1,2,\ldots,\infty$, the quantity $n^{-\frac{1}{2}}(\hat{\theta}_\psi(\underline{X}) - \theta)$ converges in distribution to a Gaussian random variable with mean zero and variance $\sigma^2(\psi,p;\rho)$ given by[3]

---

[2]Note that the sample mean can be replaced by any other consistent estimate of θ in this analysis (see [18]), in which case the condition $\sigma^2 < \infty$ might be relaxed in Theorem 1.

[3]As in [1], by $O(\rho^2)$ we mean $\lim_{\rho \to 0} |O(\rho^2)/\rho^2| < \infty$.

$$\sigma^2(\psi,p;\rho) = V(\psi,p;\rho) + O(\rho^2) \tag{8}$$

where

$$V(\psi,p;\rho) = \frac{E\{\psi^2(Y_1)\}}{(E\{\psi'(Y_1)\})^2} + 4\rho\,\frac{E\{Y_1\psi(Y_1)\}}{E\{\psi'(Y_1)\}} \tag{9}$$

with expectations computed under the assumption that $Y_1 \sim p$.

Discussion: The validity of Theorem 2 relies on two basic results: a central limit theorem (Theorem A.4 of [18]) yielding the asymptotic normality of $\hat{\theta}_\psi(\underline{X})$ and an approximation theorem (Theorem 2.1 of [18]) yielding (8). The proof of the central-limit part of this theorem relies in part on a Berry-Esseen type theorem for m-dependent random variables due to Stein [27, Corollary 3.1] which gives an error bound of $O(n^{-\frac{1}{2}})$ for the normal approximation to the distribution of $n^{-\frac{1}{2}}(\hat{\theta}_\psi(\underline{X}) - \theta)$. Thus, under the conditions of Theorem 2, we may write the false-alarm and detection probabilities of φ_ψ as

$$P_F(\varphi_\psi|p) = 1 - \Phi(n^{\frac{1}{2}}\tau/[V(\psi,p;\rho) + O(\rho^2)]^{\frac{1}{2}}) + O(n^{-\frac{1}{2}}) \tag{10}$$

and

$$P_D(\varphi_\psi|p) = 1 - \Phi(n^{\frac{1}{2}}(\tau - \theta)/[V(\psi,p;\rho) + O(\rho^2)]^{\frac{1}{2}}) + O(n^{-\frac{1}{2}}), \tag{11}$$

respectively, where Φ is the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt. \tag{12}$$

In view of (10) and (11) we adopt, for analytical purposes, the following large-sample weak-dependence approximations to P_F and P_D ,

$$\hat{P}_F(\varphi_\psi;p) = 1 - \Phi(n^{\frac{1}{2}}\tau/[V(\psi,p;\rho)]^{\frac{1}{2}}) \tag{13}$$

and

$$\hat{P}_D(\varphi_\psi;p) = 1 - \Phi(n^{\frac{1}{2}}(\tau-\theta)/[V(\psi,p;\rho)]^{\frac{1}{2}}) \ . \tag{14}$$

Note that, for a fixed noise-generating density $p \in \mathcal{F}$ and threshold $\tau \in (0,\theta)$, the error probabilities of (13) and (14) can be optimized simultaneously by minimizing the functional $V(\psi,p;\rho)$ over an appropriate class of influence curves. Note also that, for fixed p, we have

$$V(\psi,p;\rho) = [e(\psi;\rho)]^{-1} \tag{15}$$

where $e(\psi;\rho)$ is the first-order approximation developed in Part I of this study (see Lemma 1 of [1]) for the efficacy or differential signal-to-noise ratio of a detector for (1) and (2) based on comparing the detection statistic

$$\sum_{i=1}^{n} \psi(X_i) \tag{16}$$

to a threshold. Thus, the criterion of maximum $e(\psi;\rho)$ which was developed in Part I in the context of efficient detection for (1) and (2) with the structure of (16) is equally valid for approximately optimum detection in the situation with the structure of (7). In particular, by applying Theorem 1 of [1] we have that, within mild regularity on p, the problem $\min_\psi V(\psi,p;\rho)$ is solved for fixed p by the influence curve

$$\psi(x) = -p'(x)/p(x) - 2\rho I(p)x/(1+2\rho) \tag{17}$$

where p' denotes the derivative of p and where I(p) is Fisher's information number for location of p defined by

$$I(p) = \int_{-\infty}^{\infty} [(p'(x))^2/p(x)]dx \ . \tag{18}$$

Note, however, that the influence curve defined in (17) does not satisfy the hypotheses of Theorems 1 and 2 above since this function is not bounded. However, as is shown in Section IV below, the function of (17) can be approximated by a bounded influence curve that yields essentially the same performance for small $|\rho|$. This issue is discussed further in Section IV.

III. Robust Detection in Weakly Dependent Noise

The analysis of the above section indicates that the quantities of (13) and (14) can be used to approximate the probabilities of false alarm and detection, respectively, of M-detectors operating in the presence of the noise process of (2) with small $|\rho|$. Thus, in order to design M-detectors that are robust in the presence of weak dependence, we may consider the maximin formulation defined by (5) as applied to the error-probability approximations of (13) and (14).

To seek solutions to this problem we first consider the alternate minimax problem

$$\min_{\psi \in \Psi} \max_{p \in \mathcal{F}} V(\psi,p;\rho) \tag{19}$$

where the functional V is from (9) and where Ψ is an appropriate class of influence curves. We note that a saddle-point solution to (19) will be a pair $(\psi_R,p_R) \in \Psi \times \mathcal{F}$ satisfying

$$\max_{p \in \mathcal{F}} V(\psi_R,p;\rho) = V(\psi_R,p_R;\rho) = \min_{\psi \in \Psi} V(\psi,p_R;\rho), \tag{20}$$

and the existence of such a pair is equivalent to the validity of the minimax property (see, for example, Barbu and Precupanu [28])

$$\min_{\psi \in \Psi} \max_{p \in \mathcal{F}} V(\psi,p;\rho) = \max_{p \in \mathcal{F}} \min_{\psi \in \Psi} V(\psi,p;\rho) \quad . \tag{21}$$

Note that, if $(\psi_R,p_R) \in \Psi \times \mathcal{F}$ satisfies (20), then it follows from (13) and (14) that for any threshold $\tau \in (0,\theta)$ we have

$$\max_{p \in \mathcal{F}} \hat{P}_F(\varphi_{\psi_R};p) = \hat{P}_F(\varphi_{\psi_R};p_R) \tag{22}$$

and

$$\min_{\psi \in \Psi} \max_{p \in \mathcal{F}} \hat{P}_D(\varphi_\psi;p) = \hat{P}_D(\varphi_{\psi_R};p_R)$$

$$= \min_{p \in \mathcal{F}} \hat{P}_D(\varphi_{\psi_R};p) \quad . \tag{23}$$

For a particular choice α of false-alarm probability, the threshold τ_R given by

$$\tau_R = n^{-\frac{1}{2}}[V(\psi_R p_R;\rho)]^{\frac{1}{2}} \, \Phi^{-1}(1-\alpha) \tag{24}$$

will yield a value of $\hat{P}_F(\varphi_{\psi_R};p_R) = \alpha$. Thus, in view of (22) and (23), if τ_R of (24) is between 0 and θ then the pair (ψ_R,p_R) is a saddle-point solution to (5) with P_D and P_F replaced by $\hat{P}_D$ and $\hat{P}_F$, respectively. Note that the condition $\tau_R > 0$ is satisfied if $\alpha < \frac{1}{2}$, which is the only range of interest for α. However, the condition $\tau_R < \theta$ places a lower bound on the signal strength θ, relative to α and n, for which minimaxity can be achieved with $\hat{P}_D$ and $\hat{P}_F$. Further discussion of this point is included in Section VI below.

It is noteworthy that a pair solving (20) satisfies (via (10) and (11))

$$P_F(\varphi_{\psi_R};p) \le P_F(\varphi_{\psi_R};p_R) + O(\rho^2) + O(n^{-\frac{1}{2}}) \tag{25}$$

and

$$P_D(\varphi_{\psi_R};p) \ge P_D(\varphi_{\psi_R};p_R) + O(\rho^2) + O(n^{-\frac{1}{2}}) \tag{26}$$

for every $p \in \mathcal{F}$. However, corresponding $O(\rho^2)$ and $O(n^{-\frac{1}{2}})$ statements concerning $\sup_{p \in \mathcal{F}} P_F(\varphi_{\psi_R};p)$ and $\inf_{p \in \mathcal{F}} P_D(\varphi_{\psi_R};p)$ cannot be made unless the $O(\rho^2)$ and $O(n^{-\frac{1}{2}})$ terms in (8), (10), and (11) are uniform on $\mathcal{F}$. Nevertheless, because of (25) and (26), a density satisfying (20) can be considered a least-favorable member of $\mathcal{F}$ for the detection problem of interest.

In view of the above comments, we see that a solution ψ_R to (19) yields an M-detector with desirable weak-dependence large-sample robustness in performance over the class $\mathcal{F}$. The problem of (19) has been studied in [18] in the context of robust M-estimation of location. Note that, if (ψ_R, p_R) satisfies (20) and if p_R and the members of Ψ satisfy the hypothesis of Theorem 1 of [1], then Theorem 1 of [1] indicates that ψ_R must be given (up to a scale factor) by

$$\psi_R(x) = \psi_0(x) - K_R x \; ; \quad x \in \mathbb{R} \tag{27}$$

where $\psi_0 = -p_R'/p_R$ and $K_R = 2\rho I(p_R)/(1+2\rho)$, provided this function (27) is a member of Ψ. (Recall that $I(p)$ denotes Fisher's information number given by (18).) Using Eq. (36) of [1] we have (see also Eq. (3.8) of [18]) for fixed $p \in \mathcal{F}$

$$\begin{aligned} \min_{\psi} V(\psi,p;\rho) &= (1+4\rho)[I(p)]^{-1} + O(\rho^2) \\ &= (1+4\rho)\min_{\psi} V(\psi,p;0) + O(\rho^2) \quad . \end{aligned} \tag{28}$$

Thus, provided Ψ contains the appropriate influence curves, we may argue (as in [18]) that to $O(\rho^2)$ the density p_R solving $\max_{p \in \mathcal{F}} \min_{\psi \in \Psi} V(\psi,p;\rho)$

does not depend on the value of ρ. Theorem 1.1 of [18] gives a more precise reinforcement of this argument. Noting from (28) that

$$\min_{\psi} V(\psi;p;0) = [I(p)]^{-1} , \tag{29}$$

a saddle point solution to (19) thus can be sought by first choosing $p_R \in \mathcal{F}$ to solve

$$\min_{p \in \mathcal{F}} I(p) \tag{30}$$

and then choosing ψ_R from (27). Note that the resulting robust influence curve is a linearly corrected version of the influence cruve for robust M-detection in independent noise (i.e., ψ_0) as derived in [9]. This solution is thus completely analogous to the corresponding result for optimum detection with known p derived in [1].

The existence and uniqueness of solutions to (30) and their relationship to solutions to (19) with $\rho = 0$ have been studied in detail by Huber in [22]. In particular, it follows from Theorem 2 of [22] that if $\mathcal{F}$ is convex and $I(p) < \infty$ for all $p \in \mathcal{F}$, then a density p_R solving (30) and its $\rho = 0$ optimum influence curve $\psi_0 = -p_R'/p_R$ is a saddle point solution to (19) with $\rho = 0$ if $\psi_0 \in \Psi$. Furthermore, Theorem 4 of [22] implies that, if $\mathcal{F}$ is also vaguely compact, then there is a unique member of $\mathcal{F}$ solving (30). Explicit solutions to (30) for several density classes of interest and other aspects of the problem of (30) are found in Huber ([22] and [29]) and in Sullivan, et al. [30].

IV. A Modification of the Maximin Solution

As noted above, the analysis of Section IV indicates that a robust M-detector for $|\rho| \neq 0$ has an influence curve that is a linearly corrected version of the influence curve of the $\rho = 0$ robust M-detector as studied by El-Sawy and VandeLinde in [9]. This structure is undesirable for two reasons. First, as noted in Section II, an influence curve with a linear component is not bounded and thus does not satisfy the conditions needed for the validity of Theorems 1 and 2. This is only an analytical disadvantage which can be surmounted without too much difficulty. However, a more important objection to this detector is that the unboundedness of the influence curve violates basic intuitive principles of how robustness is achieved in a detector. In particular, most robust detectors provide a means for limiting the effects of extraordinarily large observations (outliers) which, if not accounted for, tend to destroy detection performance (see, for example [5]). This objection was also raised in [18] where, for the particular case in which $\mathcal{F}$ represents contaminated Gaussian noise, a truncated or lightly limited version of ψ_R is shown to produce an M-estimate which differs in worst-case performance from $\hat{\theta}_{\psi_R}$ by only $O(\rho^2)$. In this section we consider a similar modification of the robust M-detector developed in Section III in a slightly more general context.

Since the robust influence curve ψ_R of (27) is objectionable because of its unboundedness, it is reasonable (as suggested in [18]) to introduce light limiting into this structure to produce a bounded approximation to ψ_R. It is usually the case that the independent-noise robust influence curve ψ_0 is bounded (see [9] and [22]); so it is usually the linear correction term, $-K_R x$, that produces the unboundedness of ψ_R. Thus a reasonable modification of ψ_R is to replace the term $K_R x$ with a lightly limited version $K_R \ell(x)$ where $\ell(x)$ is a light limiter defined by

$$\ell(x) = \begin{cases} x & ; \text{ if } |x| \le L \\ \\ L \operatorname{sgn}(x) & ; \text{ if } |x| > L \end{cases} \tag{31}$$

with L a positive constant. Note that Chebychev's inequality implies that the probability that an individual observation exceeds the limiting point L is bounded by an upper bound proportional to L^{-2}. Thus, by choosing L sufficiently large, we should be able to make the effects of replacing x by $\ell(x)$ in ψ_R negligible. In particular, if we choose L to be $O(\rho^{-1})$, then the effects of this replacement should be $O(\rho^2)$. With this motivation, we thus propose replacing ψ_R of (27) with the modified influence curve ψ_M defined by

$$\psi_M(x) = \begin{cases} \psi_R(x) & ; \text{ if } |x| \le 1/k_\rho \\ \\ \psi_0(x) - K_R k_\rho^{-1} \operatorname{sgn}(x) & ; \text{ if } |x| > 1/k_\rho \end{cases} \tag{32}$$

where $k_\rho = O(\rho)$ and, as before, $\psi_0 = -p_R'/p_R$. For this structure we then have the following result.

<u>Theorem 3</u>: Suppose $\mathcal{F}$ is such that (30) has a solution p_R with $\psi_0 = -p_R'/p_R$ satisfying the conditions of Theorems 1 and 2. Then, for each $p \in \mathcal{F}$ such that $E\{Y_1^2\} < \infty$, we have

$$V(\psi_M, p; \rho) = V(\psi_R, p; \rho) + O(\rho^2) \quad . \tag{33}$$

Furthermore, if $E\{Y_1^2\} \leq B < \infty$ for all $p \in \mathcal{F}$, then

$$\sup_{p \in \mathcal{F}} V(\psi_M, p; \rho) = \sup_{p \in \mathcal{F}} V(\psi_R, p; \rho) + O(\rho^2) , \tag{34}$$

provided either $\rho \geq 0$ or $\rho < 0$ and $|\rho|$ is sufficiently small.

A proof of Theorem 3 is included in the appendix. As suggested above, this proof relies upon (among other things) Chebychev-type bounds on probabilities that the magnitudes of the observations exceed k_ρ^{-1}. We see that, under the conditions of this theorem, the truncated version of ψ_R yields a detector whose worst-case performance is essentially equivalent to that of φ_{ψ_R}. Note that if the conditions of Theorem 3 are not satisfied (i.e., if there is a $p \in \mathcal{F}$ such that $E\{Y_1^2\} = \infty$) then the M-detector based on ψ_R will have very poor worst-case performance relative to that of φ_{ψ_M}. Thus we may conclude generally that φ_{ψ_M} is preferable to φ_{ψ_R} from both practical and analytical viewpoints.

To illustrate the design of a robust M-detector and its modification as suggested by (32), we consider a specific example treated previously in various contexts of robust design by Huber [22], Martin and Schwartz [5], and Portnoy [18]. In particular, we consider the class $\mathcal{F}_0$ of contaminated Gaussian densities defined by

$$\mathcal{F}_0 = \{p \mid p = (1-\epsilon)\xi + \epsilon\, h;\ h \in \mathcal{H}\} \tag{35}$$

where $\xi(x) = (2\pi)^{-\frac{1}{2}}\exp\{-x^2/2\}$ is the standard Gaussian density, ϵ is a fixed number between 0 and 1, and $\mathcal{H}$ is a wide class of symmetric pdf's. Note that $\mathcal{F}_0$ thus defined can be considered to be an uncertainty neighborhood of a nominal Gaussian model with a degree ϵ of uncertainty in this model. The density minimizing Fisher's information $I(p)$ over $\mathcal{F}_0$ is given by Huber in [22] and has the well-known form

$$p_R(x) = \begin{cases} (1-\epsilon)\xi(x) & ; \ |x| \le k \\ (1-\epsilon)\xi(k)\exp\{-k|\ |x| - k|\} & ; \ |x| > k \end{cases} \tag{36}$$

where k is the unique positive solution to the equation

$$2\Phi(k) - 1 + 2\xi(k)/k = (1-\epsilon)^{-1} \quad . \tag{37}$$

The $\rho = 0$ robust influence curve is thus given by

$$\psi_0(x) = \begin{cases} x & ; \ |x| \le k \\ k \ \mathrm{sgn}(x) & ; \ |x| > k \end{cases} \tag{38}$$

and the $|\rho| \neq 0$ robust influence curve is (from (27))

$$\psi_R(x) = \begin{cases} (1-K_R)x & ; \ |x| \le k \\ k \ \mathrm{sgn}(x) - K_R x & ; \ |x| > k \quad , \end{cases} \tag{39}$$

which, for the case $\rho > 0$, increases linearly in $[-k,k]$ and decreases linearly in $[-k,k]^c$. Note that the value of $I(p_R)$ is given for this case by

$$I(p_R) = (1-\epsilon)(2\Phi(k) - 1) \ , \tag{40}$$

which, of course, must decrease monotonically with ϵ. Using (37) and (40) the values of the parameters k and K_R (recall that $K_R = 2\rho I(p_R)/(1+2\rho)$) can be computed for given ϵ and ρ. For example the case $\epsilon = 0.1$ and $\rho = 0.1$

yields $k = 1.14$ and $K_R = .115$. A reasonable bounded modification for the influence curve of (39) with $\rho > 0$ is that suggested in [18], namely

$$\psi_M(x) = \begin{cases} (1 - K_R)x & ; |x| \le k \\ k\,\mathrm{sgn}(x) - K_R x & ; k < |x| \le k_\rho^{-1} \\ 0 & ; |x| > k_\rho^{-1} \end{cases} \tag{41}$$

where $k_\rho = K_R/k$ which is $O(\rho)$ since K_R is $O(\rho)$. Figure 1 illustrates this function for the case $\epsilon = 0.1$ and $\rho = 0.1$. Further discussion of this and related examples is included in Sections V and VI below.

V. Performance of the Proposed Robust System Relative to the Independent-Noise Robust System

In Section III we argued that the least-favorable noise-generating density p_R is independent of the value of the dependence parameter ρ. Thus, for each value of ρ, the robust influence curve ψ_R is the optimum influence curve (i.e., the solution to $\min_{\psi} V(\psi,p_R;\rho)$) corresponding to the fixed density p_R. It follows from the analysis in Part I of this study (in particular, from Theorem 2 of [1]) that, within regularity conditions, the quantities $V(\psi_R,p_R;\rho)$ and $V(\psi_0,p_R;\rho)$ differ only by $O(\rho^2)$ terms where, as before, $\psi_0 = -p_R'/p_R$. Since ψ_0 is the influence curve one would use for robustness if there were no dependence, a question arises as to whether the worst-case performance[4] of φ_{ψ_0} over $\mathcal{F}$ might be only $O(\rho^2)$ different from that of φ_{ψ_R}. That this is not generally the case is shown in this section; in particular, we demonstrate that this difference in worst-case performance is actually $O(\rho)$ for most uncertainty models of interest.

To consider the worst-case performance of φ_{ψ_R} relative to the worst-case performance of φ_{ψ_0} for fixed ρ we first give the following result.

<u>Lemma 1</u>: Suppose $p_R \in \mathcal{F}$ solves (30) with $0 < I(p_R) < \infty$. Suppose further that there are numbers b and B such that $0 < b \le E\{\psi_R^2(Y_1)\} \le B < \infty$, $E\{Y_1^2\} \le B$, and $b \le E\{\psi_0'\} \le B$, for all $p \in \mathcal{F}$ where $\psi_0 = -p_R'/p_R$. Then, if $\rho \le 0$ or if $\rho > 0$ and is sufficiently small, we have

$$\frac{V(\psi_R,p;\rho)}{V(\psi_0,p;\rho)} = 1 + C_p\rho + O(\rho^2) , \qquad (42)$$

---

[4] It should be noted here that $V(\psi_0,p_R;0)$ is the worst-case value of $V(\psi_0,p;0)$ over $p \in \mathcal{F}$, but it is not necessarily true that $V(\psi_0,p_R;\rho)$ is the worst-case value of $V(\psi_0,p;\rho)$ over $\mathcal{F}$ for $\rho \ne 0$. In fact, this latter situation is <u>not</u> usually the case, as follows from Lemma 1 and Theorem 4.

where

$$C_p = \frac{4I(p_R)V(\psi_0,p;-\frac{1}{2})}{E\{\psi_0'(Y_1)\}V(\psi_0,p;0)} , \tag{43}$$

and where $O(\rho^2)$ is uniform over $\mathcal{F}$.

The proof of Lemma 1 is straightforward and will be omitted. This result shows that, although $V(\psi_R,p_R;\rho)$ and $V(\psi_0,p_R;\rho)$ differ by only $O(\rho^2)$ terms, $V(\psi_R,p;\rho)$ and $V(\psi_0,p;\rho)$ may possibly differ by $O(\rho)$ terms if $p \neq p_R$. (Note: it follows straightforwardly from (9) that $V(\psi_0,p_R;-\frac{1}{2}) = 0$.) It follows from Lemma 1 that, for fixed ρ, we have

$$\sup_{p \in \mathcal{F}} V(\psi_0,p;\rho) = \sup_{p \in \mathcal{F}} V(\psi_R,p;\rho) + C|\rho| + O(\rho^2) \tag{44}$$

where $C = \sup_{p \in \mathcal{F}} C_p$ if $\rho > 0$ and $C = |\inf_{p \in \mathcal{F}} C_p|$ if $\rho < 0$. We note, in particular, that (44) can <u>not</u> hold for any $C < 0$ since we must have

$$\begin{aligned} \sup_{p \in \mathcal{F}} V(\psi_0,p;\rho) &\geq V(\psi_0,p_R;\rho) = V(\psi_R,p_R;\rho) + O(\rho^2) \\ &= \sup_{p \in \mathcal{F}} V(\psi_R,p;\rho) + O(\rho^2) \end{aligned} \tag{45}$$

where the first equality in (45) follows from Theorem 2 of [1]. Since (45) is valid for both positive and negative ρ, (45) and (42) imply that either $V(\psi_0,p;-\frac{1}{2})$ takes on both negative and positive values over $\mathcal{F}$ or $V(\psi_0,p;-\frac{1}{2})$ is identically zero on $\mathcal{F}$. Thus, for every ρ there is a $C \geq 0$ such that (44) holds, and $C = 0$ if, and only if, $V(\psi_0,p;-\frac{1}{2}) = 0$ for all $p \in \mathcal{F}$. We note again that C depends only on the algebraic sign of ρ. Some conditions under which $C \neq 0$ are summarized in the following result.

<u>Theorem 4</u>: Suppose $\mathcal{F}$ is such that the following three conditions hold:

(i) There is a set $\Omega \subset \mathbb{R}$ and a constant $K > 0$ such that $x \in \Omega$ implies $\psi_0(x) = K\, \mathrm{sgn}(x)$.

(ii) There exist pdf's p_1 and p_2 in $\mathcal{F}$ such that $p_1 = p_2 = p_R$

on Ω^c and $\int_{-\infty}^{\infty} |y| p_1(y)dy < \int_{-\infty}^{\infty} |y| p_R(y)dy < \int_{-\infty}^{\infty} |y| p_2(y)dy$.

(iii) $0 < I(p_R) < \infty$ and $\lim_{|y| \to \infty} \psi_0(y)p'(y) = 0$ for $p = p_R$, $p = p_1$, and $p = p_2$.

Then $V(\psi_0, p_1; -\frac{1}{2}) > 0$ and $V(\psi_0, p_2; -\frac{1}{2}) < 0$, and hence the constant C of (44) is positive.

<u>Proof</u>: See the appendix.

Thus we see that, under the conditions of Lemma 1 and Theorem 4, the improvement in worst-case performance by using ψ_R rather than ψ_0 is of first order in $|\rho|$. Theorem 3 implies that ψ_M of (32) also yields this $O(\rho)$ improvement. Note that the conditions of Theorem 4 are satisfied by most of the usual models for distributional uncertainty. For example, the contaminated-Gaussian class treated in Section IV satisfies Conditions (i) through (iii) as is easily seen from (35), (36), (38), and (40). Other classes that satisfy these conditions include contaminated-mixture classes with nominal models other than Gaussian (see Huber [22] and Kassam and Thomas [8]), p-point classes as considered by El-Sawy and VandeLinde [9,10], and the class of densities whose cumulative probability distribution functions differ in sup-norm from the standard Gaussian by no more than some prescribed amount (see [22]). Thus, we may conclude that ψ_M is generally preferable to ψ_0 to the extent that $O(\rho)$ terms are appreciable (i.e., to the extent that the model of (2) is of interest).

VI. Summary and Discussion

In this paper we have considered the problem of designing robust systems for the detection of signals in weakly dependent noise. To find solutions to this problem we have considered the class of M-detectors that was proposed by El-Sawy and VandeLinde in [9] for robustness in the corresponding independent-noise case. For this class of detectors it was seen that the robust estimation analysis of Portnoy [18] is applicable to the design of robust structures for the weakly dependent moving-average model of noise dependence. In particular, it was seen that an α-level robust M-detector φ_{ψ_R} for a class $\mathcal{F}$ of noise-generating p.d.f.'s can be sought by choosing p_R to minimize Fisher's information $I(p)$ over $\mathcal{F}$ and then taking ψ_R from (27). The detection threshold τ_R is chosen by (24) and must lie between 0 and θ for approximate minimaxity. Since ψ_R is usually not bounded it is intuitively more reasonable to modify ψ_R by truncating the linear term at points $\pm k_\rho^{-1}$ where k_ρ is $O(\rho)$ as in (32). The resulting M-detector φ_{ψ_M} is equivalent to φ_{ψ_R} to $O(\rho^2)$ under the conditions of Theorem 3. The worst-case performance of both of these detectors is better by $O(\rho)$ than that of the $\rho = 0$ robust M-detector φ_{ψ_0} under the conditions of Lemma 1 and Theorem 4. As is the case with results of Part I of this study [1], the results of this paper can also be extended straightforwardly to moving averages of higher order than (2) by applying the results of Portnoy [19] for M-estimation in such models. However, the basic structure and performance of the robust M-detector are unchanged by this generalization.

In general, to implement the robust detector φ_{ψ_R} one must first compute $\hat{\theta}_{\psi_R}(\underline{x})$ from (7) and then compare this value to a threshold. However, note that, if $\psi_R(x)$ is a strictly increasing function of x, then $\sum_{i=1}^{n} \psi_R(x_i - T)$ is a strictly decreasing function of T; and, in this case, φ_{ψ_R} can be implemented as follows (see [9]):

$$\varphi_{\psi_R}(\underline{x}) = \begin{cases} 1 \ ; & \text{if } \sum_{i=1}^{n} \psi_R(x_i - T_\tau) < 0 \\ \gamma \ ; & \text{if } \sum_{i=1}^{n} \psi_R(x_i - T_\tau) = 0 \\ 0 \ ; & \text{if } \sum_{i=1}^{n} \psi_R(x_i - T_\tau) > 0 \end{cases} \tag{46}$$

where T_τ and γ are chosen to give desired false-alarm performance. The structure of (46) is simpler to implement than is (6) since it is not necessary to solve (7) to perform the test of (46). However, it is not always the case that ψ_R is increasing, although it will be increasing if

$$d\psi_0(x)/dx > K_R \text{ for all } x \in \mathbb{R} \ . \tag{47}$$

If p_R is strongly unimodal (i.e., $-\log(p_R)$ is convex) then (47) holds for all $\rho < 0$; however, for $\rho > 0$, (47) does not hold for many practical cases even when p_R is strongly unimodal because of the redescending nature of ψ_R (such as in the contaminated-Gaussian example of Section IV). If ψ_R is not strictly increasing, then (46) cannot be used and (7) must be solved; however, efficient iterative techniques for solving (7) have been developed (see, for example, Collins [31]).

The robustness of the proposed M-detector is restricted to situations for which the threshold τ_R of (24) is between 0 and θ. As noted above, the trivial condition $\alpha < \frac{1}{2}$ is sufficient for τ_R to be positive; however, the condition that τ_R be less than θ places a lower bound on θ, for fixed α and n, for which φ_{ψ_R} is a maximin solution. If θ is smaller than is necessary the alternate approach of local robustness (see, for example, Kassam and Thomas [8]) can possibly be applied. However, the redescending nature of the solution to $\min_{\psi} V(\psi,p;\rho)$ for $\rho > 0$ may cause problems in the local formulation (see also Kassam, et al. [32]).

Before concluding, two comments concerning the overall optimality of the proposed robust detector are in order. First, since we have considered only M-detectors, a question arises as to the performance of the proposed detector relative to other detectors that are not of this form. In answer to this question we note that it has been demonstrated in [18] (Theorem 1.1) that, within regularity, the optimum M-estimate of θ in (1) has variance $O(\rho^2)$ close to the theoretical minimum possible variance for asymptotically Gaussian unbiased estimates of θ. Thus, among threshold detectors based on estimates of θ, the class of M-detectors are capable of achieving overall optimum performance to $O(\rho^2)$. As a second comment we note that higher order (in ρ) expressions for the asymptotic variance of $\hat{\theta}_{\psi}(\underline{X})$ are available (see [18, p. 39] and [19, Lemma 2.1]); thus the question of whether or not a corresponding higher order optimum influence curve is possible arises. However, that no such M-estimate exists follows from Theorem 2.1 of [19] which states that, for fixed p and within mild assumptions, no M-estimate depending only on ρ can achieve variance closer than $O(\rho^2)$ to the theoretical minimum variance. This implies a similar statement for M-detectors.

Appendix

A. A Proof of Theorem 3

Define $Z = \max\{-k_\rho^{-1}, \min\{Y_1, k_\rho^{-1}\}\}$. We have

$$E\{\psi_M^2(Y_1)\} = E\{\psi_R^2(Y_1)\} + 2E\{(\psi_M(Y_1) - \psi_R(Y_1))\psi_R(Y_1)\}$$

$$+ E\{|\psi_M(Y_1) - \psi_R(Y_1)|^2\}$$

$$= E\{\psi_R^2(Y_1)\} + 2K_R E\{(Z - Y_1)\psi_0(Y_1)\}$$

$$- K_R^2\, E\{(Z - Y_1)Y_1\} + K_R^2\, E\{|Z - Y_1|^2\} \,. \tag{A1}$$

Note that $|Z - Y_1| \leq |Y_1|$ so that the last two terms in the right-hand side of (A1) are each bounded in magnitude by $K_R^2\, E\{Y_1^2\}$. Concerning the second term in the right-hand side of (A1) we have (applying the Schwarz and Chebychev inequalities)

$$|E\{(Z - Y_1)\psi_0(Y_1)\}|^2 \leq E\{(Z - Y_1)^2\psi_0^2(Y_1)\}P\{|Z - Y_1| > 0\}$$

$$\leq E\{Y_1^2\} \sup_{y \in \mathbb{R}} \psi_0^2(y) P\{|Y_1| > k_\rho^{-1}\}$$

$$\leq E^2\{Y_1^2\} \sup_{y \in \mathbb{R}} \psi_0^2(y) k_\rho^2 \,.$$

Thus, since k_ρ and K_R are $O(\rho)$, we have

$$E\{\psi_M^2(Y_1)\} = E\{\psi_R^2(Y_1)\} + O(\rho^2) \,. \tag{A2}$$

Similarly, we have

$$E\{\psi_M'(Y_1)\} = E\{\psi_R'(Y_1)\} + K_R P\{|Y_1| > k_\rho^{-1}\}$$

$$= E\{\psi_R'(Y_1)\} + K_R O(\rho^2) \tag{A3}$$

and

$$E\{Y_1\psi_M(Y_1)\} = E\{Y_1\psi_R(Y_1)\} + K_R E\{(Y_1 - Z)\}$$

$$= E\{Y_1\psi_R(Y_1)\} + O(\rho^2) \; . \tag{A4}$$

Equations (A2) through (A4) imply (33). In each of (A2), (A3), and (A4) the higher order terms in ρ are uniform over $\mathcal{F}$ if $E\{Y_1^2\}$ has a uniform bound. Thus, in this case, the $O(\rho^2)$ in (33) is uniform if (A3) is nonzero over $\mathcal{F}$, a condition which occurs if $\inf\limits_{p \in \mathcal{F}} E\{\psi_R'(Y_1)\} > -K_R B\, k_\rho^2$. The rest of Theorem 3 follows.

B. <u>A Proof of Theorem 4</u>

The sign of $V(\psi_0,p;-\frac{1}{2})$ is the same as the sign of

$$\int \psi_0^2\, p - (\int \psi_0'\, p)(\int_{-\infty}^{\infty} y\psi_0(y)p(y)dy). \tag{B1}$$

Writing $p = p_R + (p - p_R)$, we see that the quantity of (B1) is equal to

$$\int \psi_0^2 p_R - (\int\psi_0' p_R)(\int_{-\infty}^{\infty} y\psi_0(y)p_R(y)dy) + \int\psi_0^2(p - p_R)$$

$$- (\int\psi_0'(p - p_R))(\int_{-\infty}^{\infty} y\psi_0(y)p_R(y)dy) - (\int\psi_0' p_R)(\int_{-\infty}^{\infty} y\psi_0(y)(p(y) - p_R(y))dy. \tag{B2}$$

The first two terms of (B2) add to zero. If p and p_R differ only on Ω the third and fourth terms of (B2) are both zero and the final term of (B2) becomes

$$I(p_R)K[\int_{-\infty}^{\infty} |y|\, p_R(y)dy - \int_{-\infty}^{\infty} |y|\, p(y)dy] \; . \tag{B3}$$

Theorem 4 follows from (B3).

References

[1] H. V. Poor, "Signal detection in the presence of weakly dependent noise - Part I: Optimum detection," IEEE Trans. Inform. Theory, vol. IT-28, 1982 (to appear).

[2] P. J. Huber, "A robust version of the probability ratio test," Ann. Math. Stat., vol. 36, pp. 1753-1758, 1965.

[3] P. J. Huber and V. Strassen, "Minimax tests and the Neyman-Pearson lemma for capacities," Ann. Statist., vol. 1, pp. 251-263, 1973.

[4] H. V. Poor, "Robust decision design using a distance criterion," IEEE Trans. Inform. Theory, vol. IT-26, pp. 575-587, September 1980.

[5] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," IEEE Trans. Inform. Theory, vol. IT-17, pp. 50-56, January 1971.

[6] J. Capon, "On the asymptotic efficiency of locally optimum tests," IRE Trans. Inform. Theory, vol. IT-7, pp. 67-71, April 1961.

[7] T. S. Ferguson, Mathematical Statistics: A Decision Theoretic Approach. New York: Academic, 1967.

[8] S. A. Kassam and J. B. Thomas, "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," IEEE Trans. Inform. Theory, vol. IT-22, pp. 22-26, January 1976.

[9] A. H. El-Sawy and V. D. VandeLinde, "Robust detection of known signals," IEEE Trans. Inform. Theory, vol. IT-23, pp. 722-727, November 1977.

[10] A. H. El-Sawy and V. D. VandeLinde, "Robust sequential detection of signals in noise," IEEE Trans. Inform. Theory, vol. IT-25, pp. 346-353, May 1979.

[11] M. Kanefsky and J. B. Thomas, "On adaptive nonparametric detection systems using dependent samples," IEEE Trans. Inform. Theory, vol. IT-11, pp. 521-526, October 1965.

[12] J. L. Gastwirth, H. Rubin, and S. S. Wolff, "The effect of autoregressive dependence on a non-parametric test," IEEE Trans. Inform. Theory, vol. IT-13, pp. 311-313, April 1967.

[13] L. D. Davisson, E. A. Feustal, and J. W. Modestino, "The effects of dependence on nonparametric detection," IEEE Trans. Inform. Theory, vol. IT-16, pp. 32-41, January 1970.

[14] S. A. Kassam and J. B. Thomas, "A class of nonparametric detectors for dependent input data," IEEE Trans. Inform. Theory, vol. IT-21, pp. 431-437, July 1975.

[15] M. B. Sirvanci and S. S. Wolff, "Nonparametric detection with autoregressive data," IEEE Trans. Inform. Theory, vol. IT-22, pp. 725-731, November 1976.

[16] A. Hoyland, "Robustness of the Wilcoxon estimate of location against a certain dependence," Ann. Math. Stat., vol. 39, pp. 1196-1201, 1968.

[17] J. S. Gastwirth and H. Rubin, "The behavior of robust estimators on dependent data," Ann. Statist., vol. 3, pp. 1070-1100, 1975.

[18] S. L. Portnoy, "Robust estimation in dependent situations," Ann. Statist., vol. 5, pp. 22-43, 1977.

[19] S. L. Portnoy, "Further remarks on robust estimation in dependent situations," Ann. Statist., vol. 7, pp. 224-231, 1979.

[20] H. L. Koul, "Behavior of robust estimators in the regression model with dependent errors," Ann. Statist., vol. 5, pp. 681-699, 1977.

[21] R. D. Martin and G. M. deMontricher, "Locally max-min robust tests," Technical Report No. 61, Series 2, Princeton University Department of Statistics, Princeton, NJ, May 1974.

[22] P. J. Huber, "Robust estimation of a location parameter," Ann. Math. Stat., vol. 35, pp. 73-101, 1964.

[23] F. R. Hampel, "A general qualitative definition of robustness," Ann. Math. Stat., vol. 42, pp. 1887-1896, 1971.

[24] P. J. Huber, Robust Statistical Procedures. Philadelphia: SIAM, 1977.

[25] S. D. Silvey, Statistical Inference. London: Chapman and Hall, 1975.

[26] P. Billingsley, Convergence of Probability Measures. New York: Wiley, 1968.

[27] C. Stein, "A bound on the error in the normal approximation to the distribution of the sum of dependent random variables," in Proc. Sixth Berkeley Symp. Math. Statist. Prob. - Vol. II, Probability Theory, L. M. Le Cam et al., eds. Berkeley, California: University of California Press, 1972, pp. 583-602.

[28] V. Barbu and Th. Precupanu, Convexity and Optimization in Banach Spaces. Bucharest: Editura Academici, 1978.

[29] P. J. Huber, "Fisher information and spline interpolation," Ann. Statist., vol. 2, pp. 1029-1033, 1974.

[30] F. J. M. Sullivan, V. D. VandeLinde, and S. S. Wolff, "Densities with minimum Fisher information for location," Technical Report No. JHU-EE73-10, The Johns Hopkins University Department of Electrical Engineering, June 1973.

[31] J. R. Collins, "Robust estimation of a location parameter in the presence of asymmetry," Ann. Statist., vol. 4, pp. 68-85, 1976.

[32] S. A. Kassam, G. Moustakides, and J. G. Shin, "Robust detection of known signals in asymmetric noise," IEEE Trans. Inform. Theory, vol. IT-27, 1981 (to appear).

Footnotes

1. Recall that the local power slope is the criterion for designing locally most powerful detection systems (see Capon [6] or Ferguson [7]).

2. Note that the sample mean can be replaced by any other consistent estimate of θ in this analysis (see [18]), in which case the condition $\sigma^2 < \infty$ might be relaxed in Theorem 1.

3. As in [1], by $O(\rho^2)$ we mean $\lim_{\rho \to 0} |O(\rho^2)/\rho^2| < \infty$.

4. It should be noted here that $V(\psi_0, p_R; 0)$ is the worst-case value of $V(\psi_0, p; 0)$ over $p \in \mathcal{F}$, but it is not necessarily true that $V(\psi_0, p_R; \rho)$ is the worst-case value of $V(\psi_0, p; \rho)$ over $\mathcal{F}$ for $\rho \neq 0$. In fact, this latter situation is not usually the case, as follows from Lemma 1 and Theorem 4.

Figure Caption

Fig. 1 - Influence curve for robust M-detection in dependent contaminated Gaussian noise with $\epsilon = 0.1$ and $\rho = 0.1$.

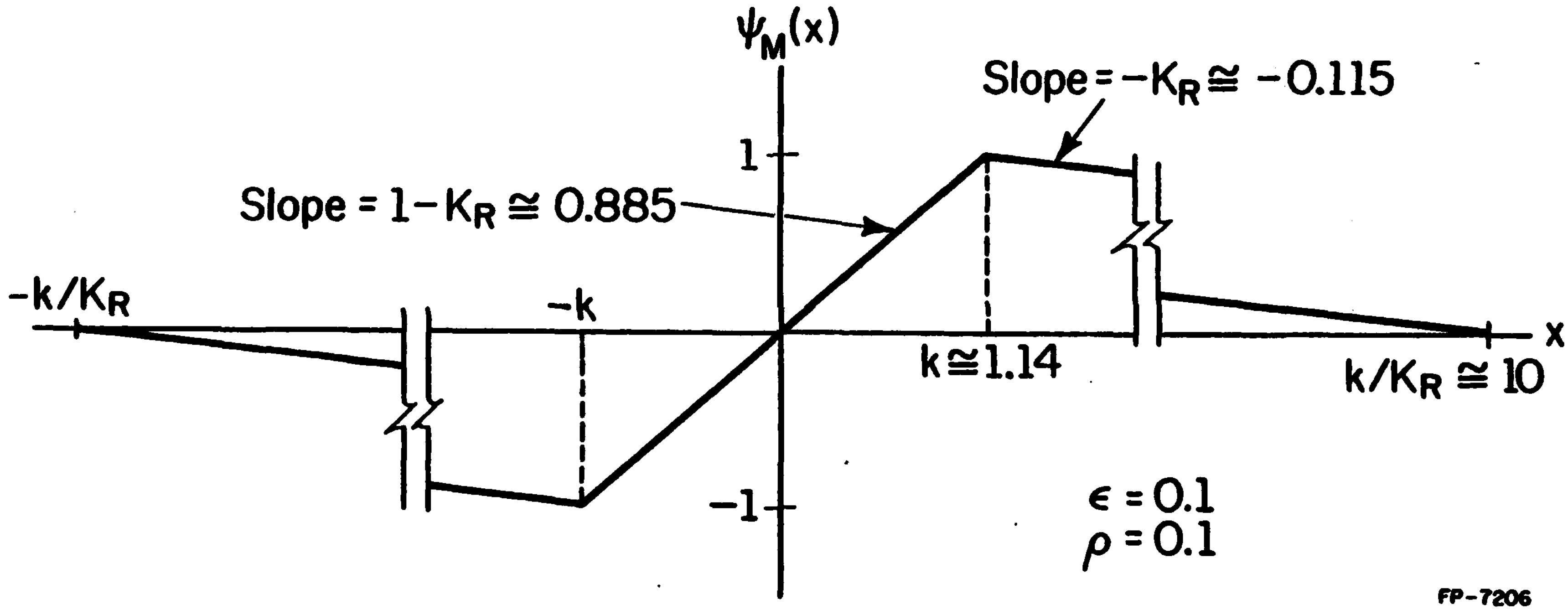


Figure 1. Influence curve for robust M-detection in dependent contaminated Gaussian noise with $\epsilon = 0.1$ and $\rho = 0.1$.

END